Oxford, 2 February 2017

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Climate prediction: Climate modelling driven by climate services

Francisco J. Doblas-Reyes
BSC Earth Sciences Department

iCrea

BSC Earth Sciences Department

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

What

Environmental forecasting

Why

Our strength …
- … research …
- … operations …
- … services …
- … high resolution …

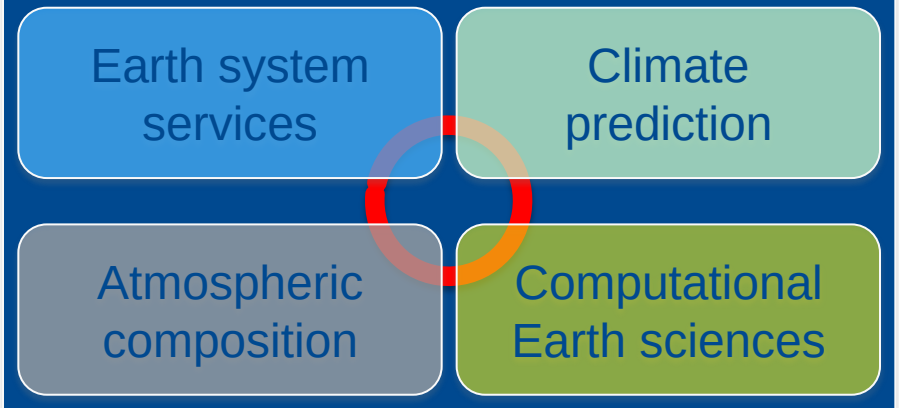
How

Develop a capability to model air quality processes from urban to global and the impacts on weather, health and ecosystems

Implement a climate prediction system for subseasonal-to-decadal climate prediction

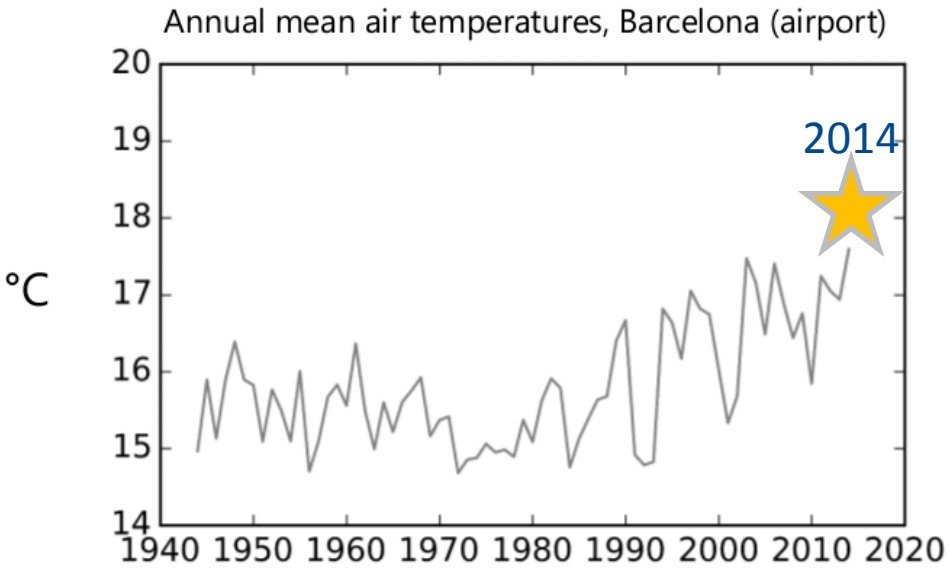
Develop user-oriented services that favour both technology transfer and adaptation

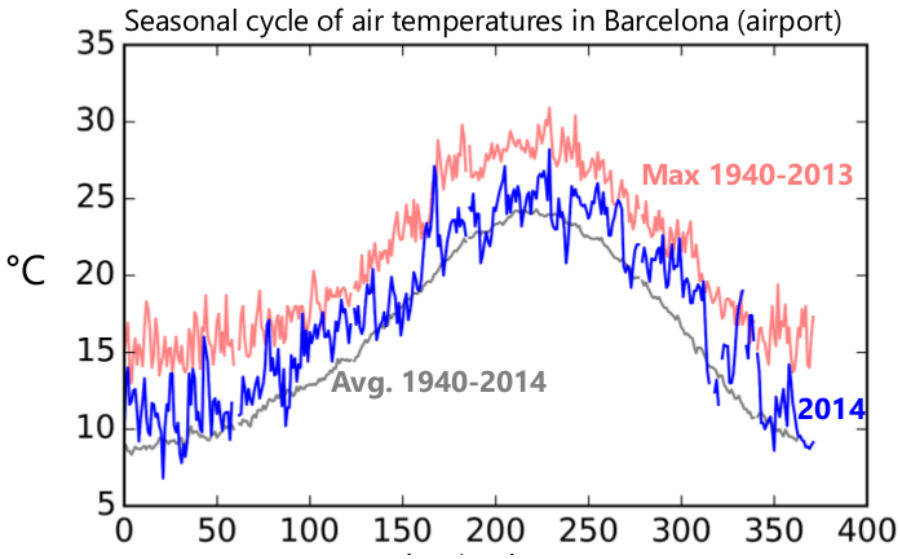
Use cutting-edge HPC and Big Data technologies for the efficiency and user-friendliness of Earth system models



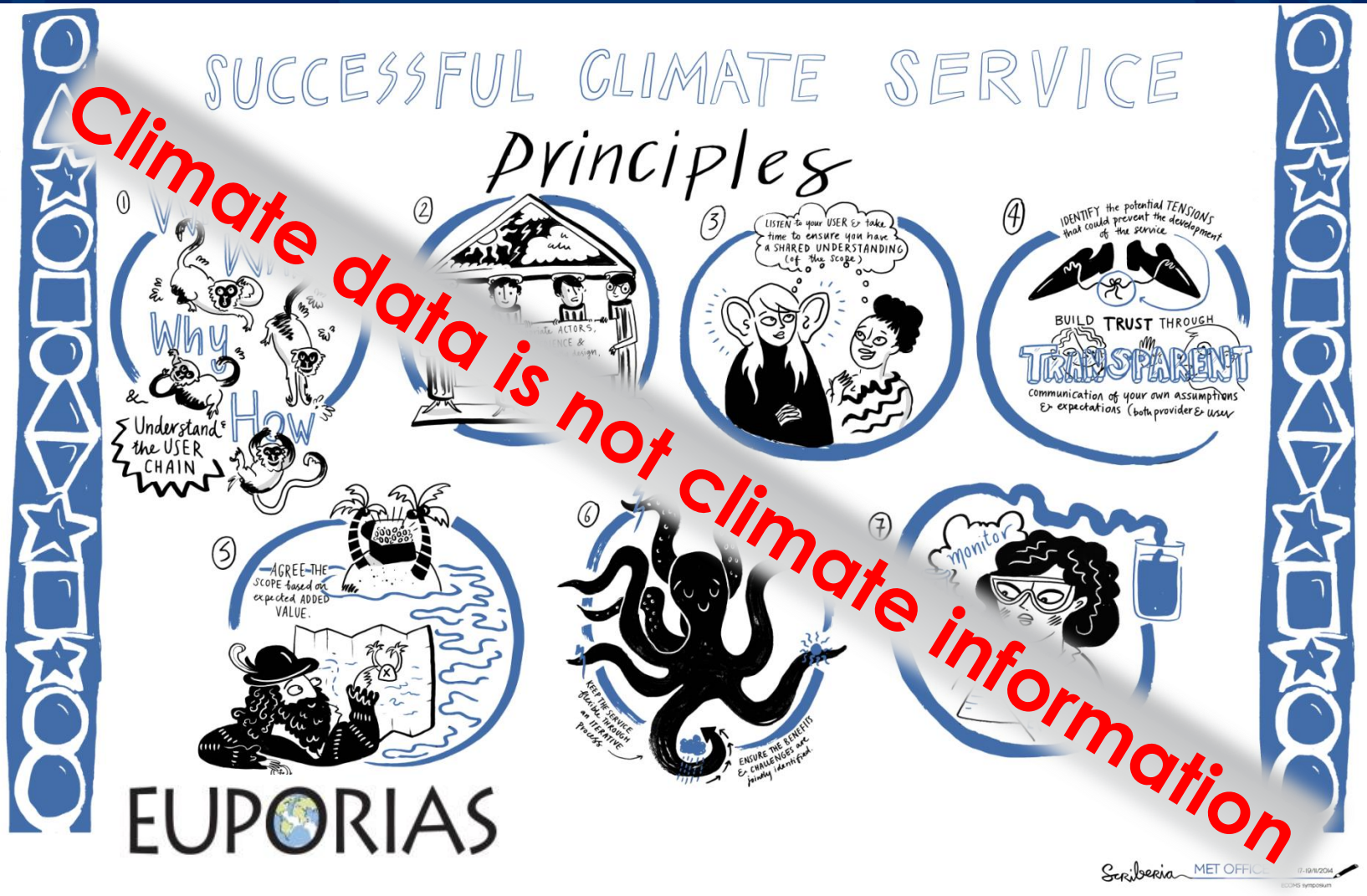
2014, a special year

Temperatures in Barcelona airport from the ECAD dataset.



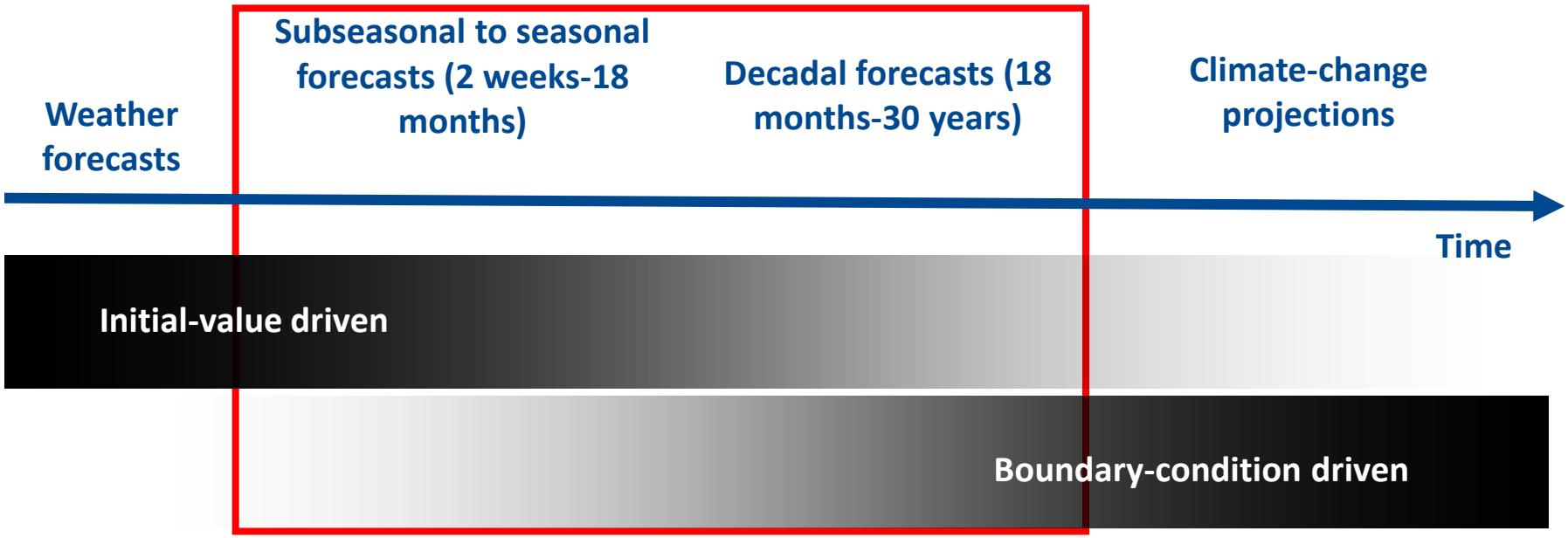


Service-driven climate research



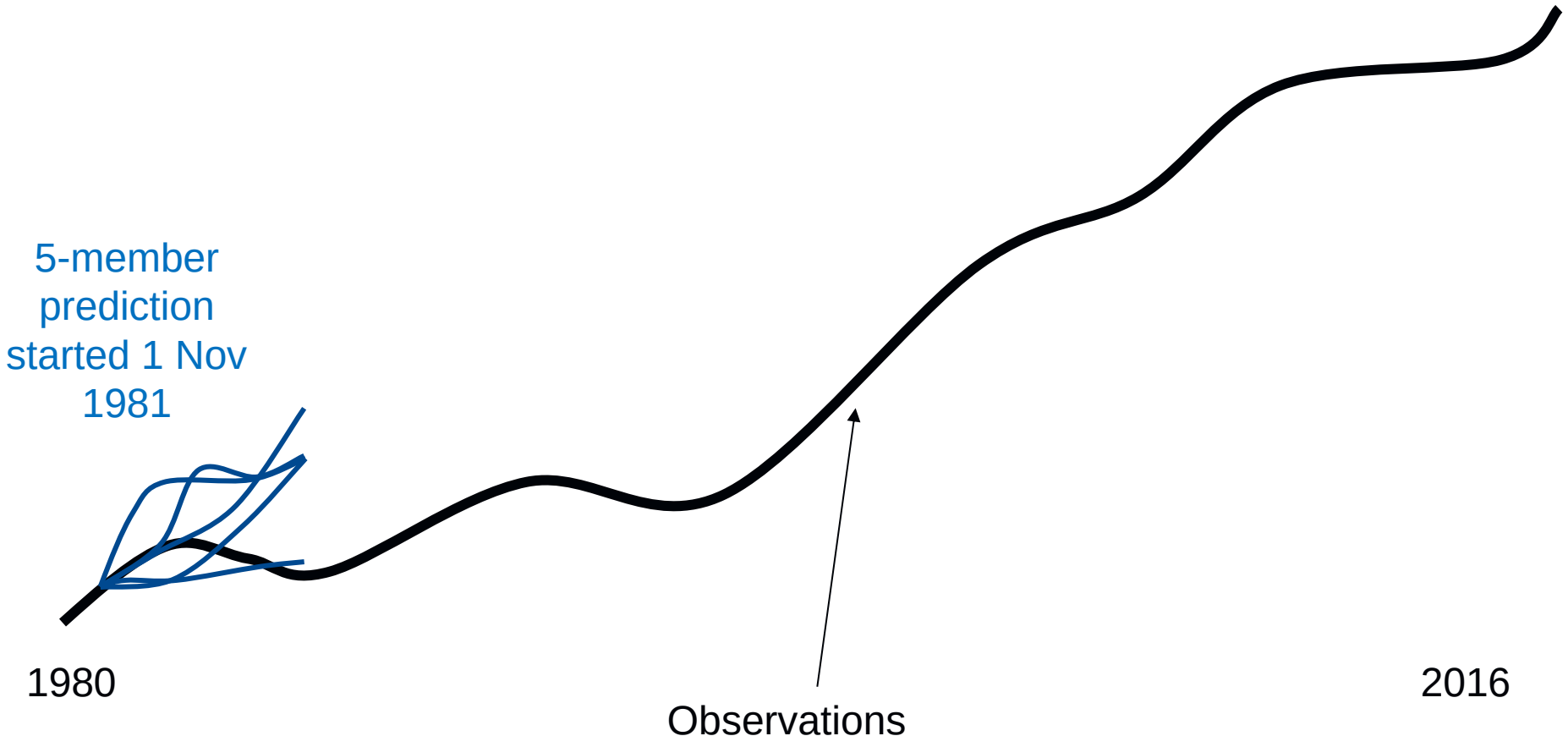
Climate prediction time scales

Progression from initial-value problems with weather forecasting at one end and multi-decadal to century projections as a forced boundary condition problem at the other, with climate prediction (sub-seasonal, seasonal and decadal) in the middle. Prediction involves initialization and systematic comparison with a simultaneous reference.

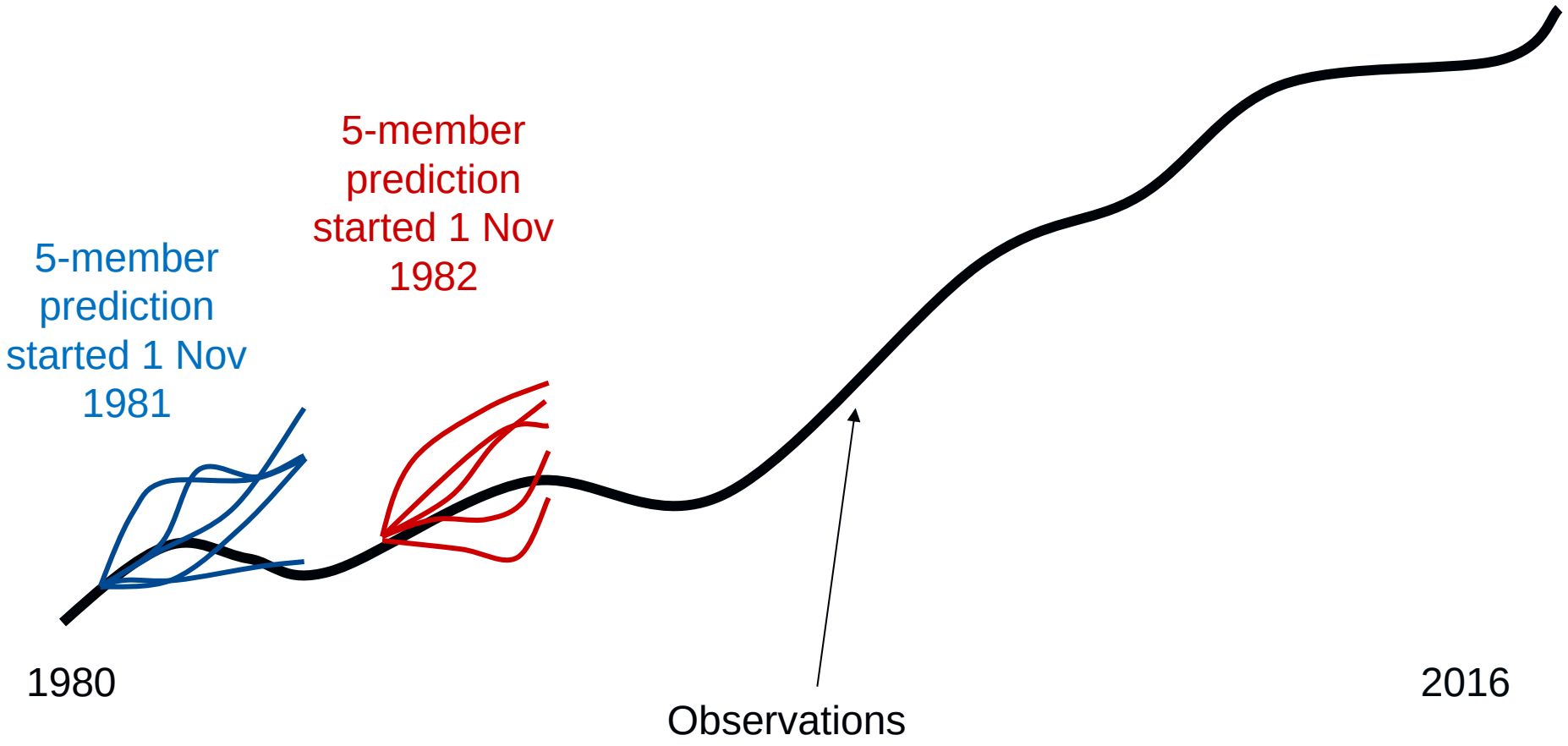


Adapted from Meehl et al. (2009)

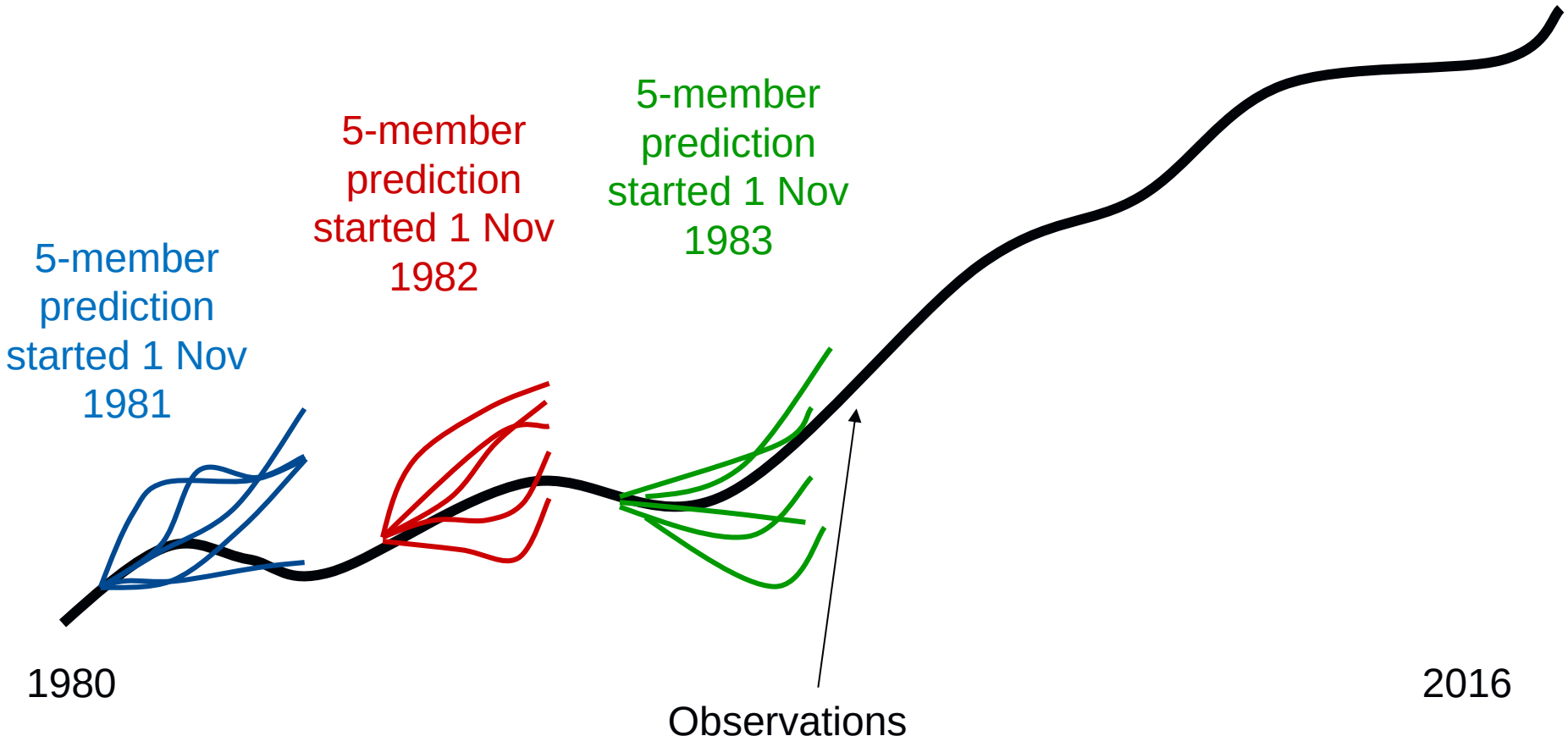
Climate prediction experiments



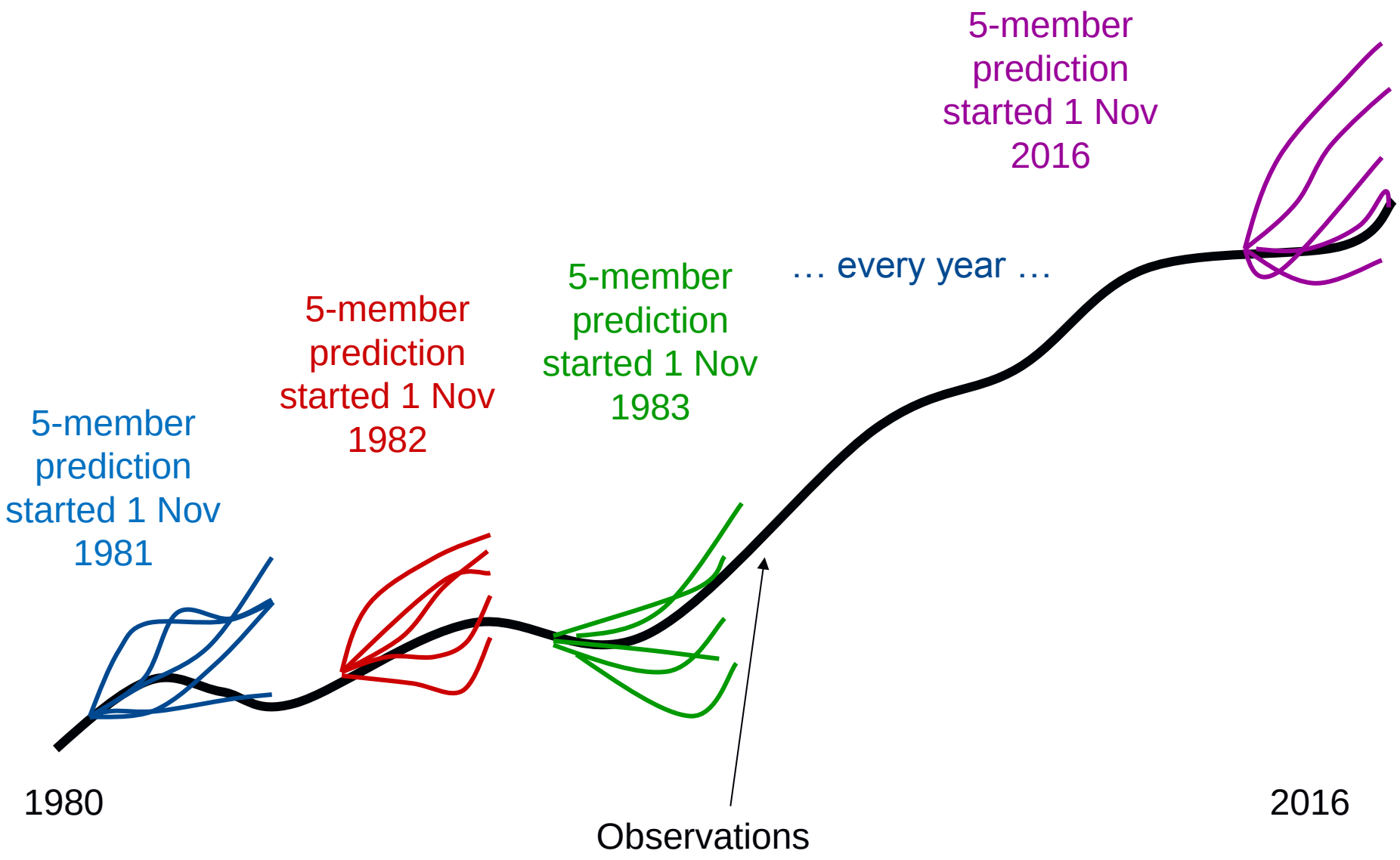
Climate prediction experiments



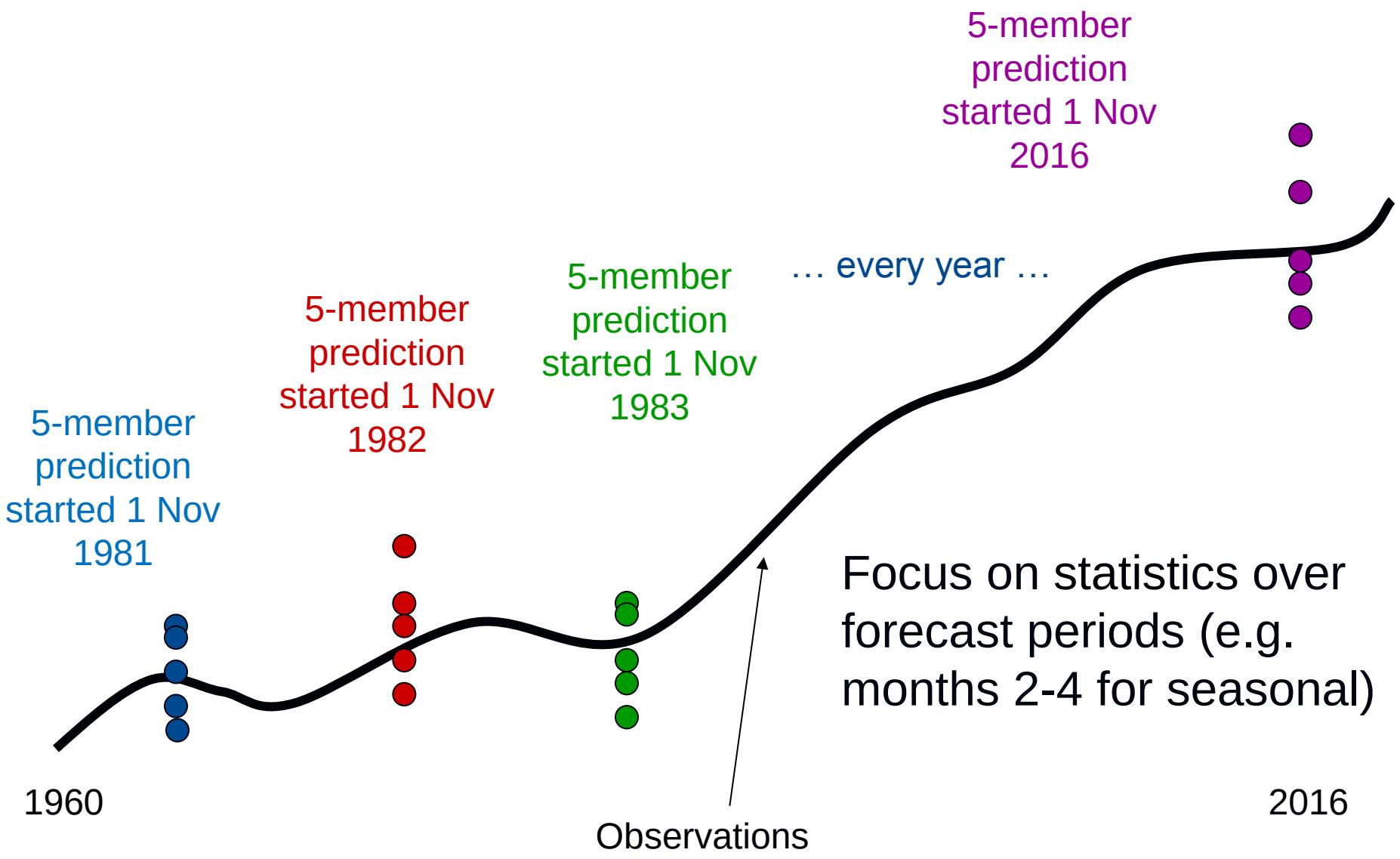
Climate prediction experiments



Climate prediction experiments

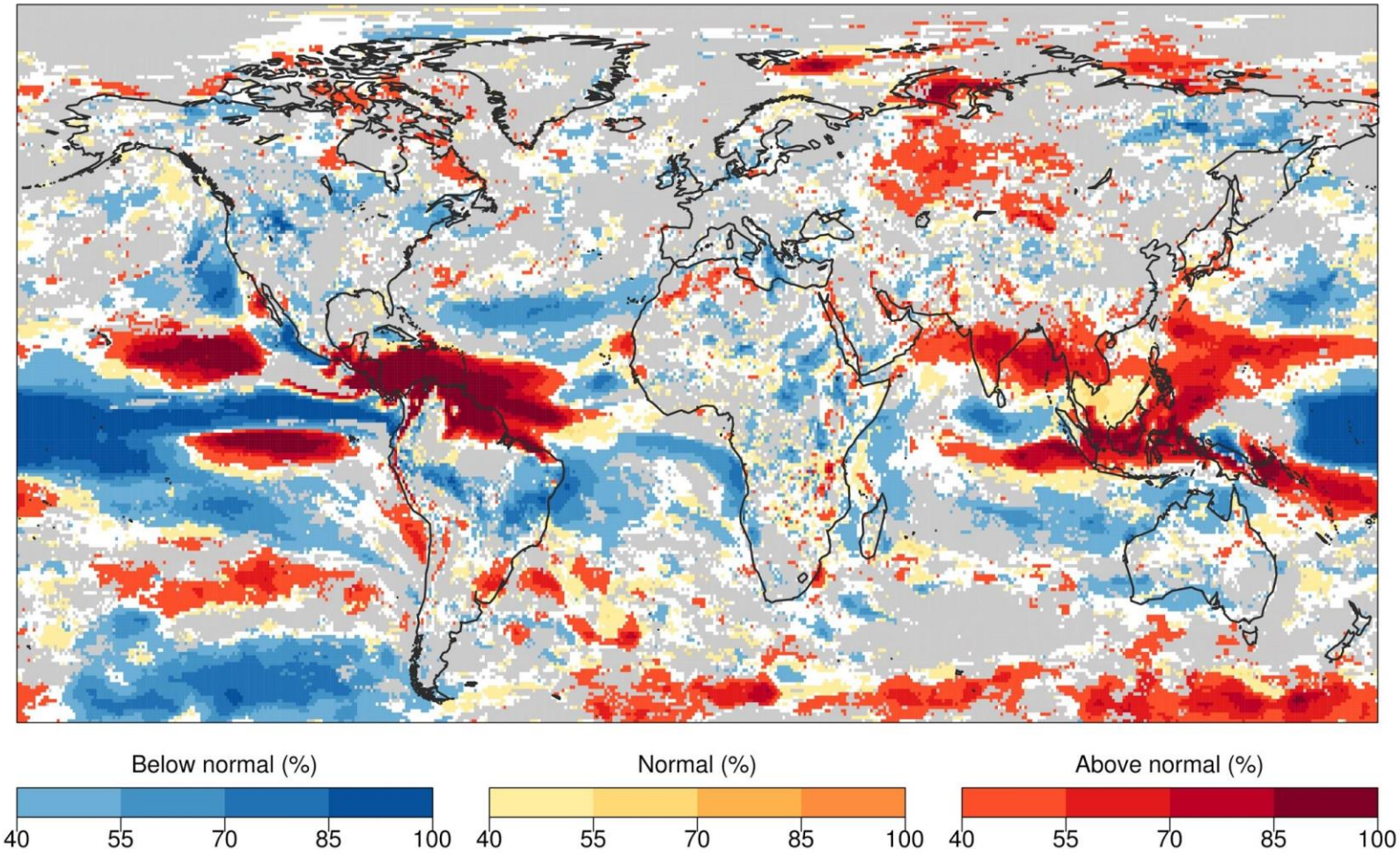


Climate prediction experiments



Focus on statistics over forecast periods (e.g. months 2-4 for seasonal)

Displaying predictions



Wind speed prediction for June 1st - August 31st 2015, issued on May 1st 2005.

The most likely wind power category (below normal, normal or above normal), and its percentage probability to occur is shown. "Normal" represents the average of the past. White areas show where the probability is <40% and approximately equal for all three categories. Grey areas show where the climate prediction model does not improve upon the standard and current approach, which projects past climate data into the future.

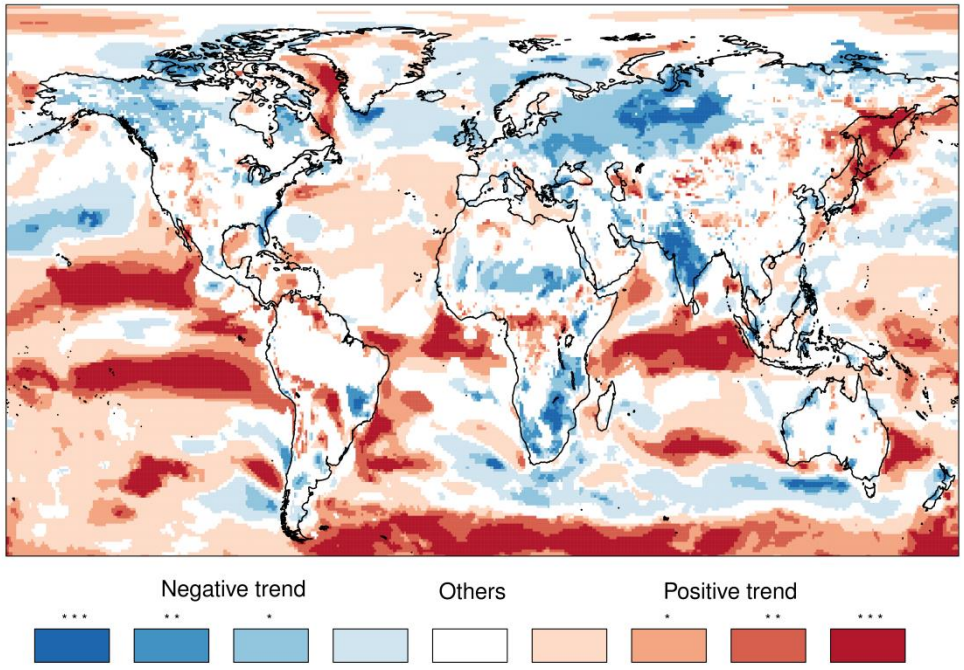
Some problems of this product

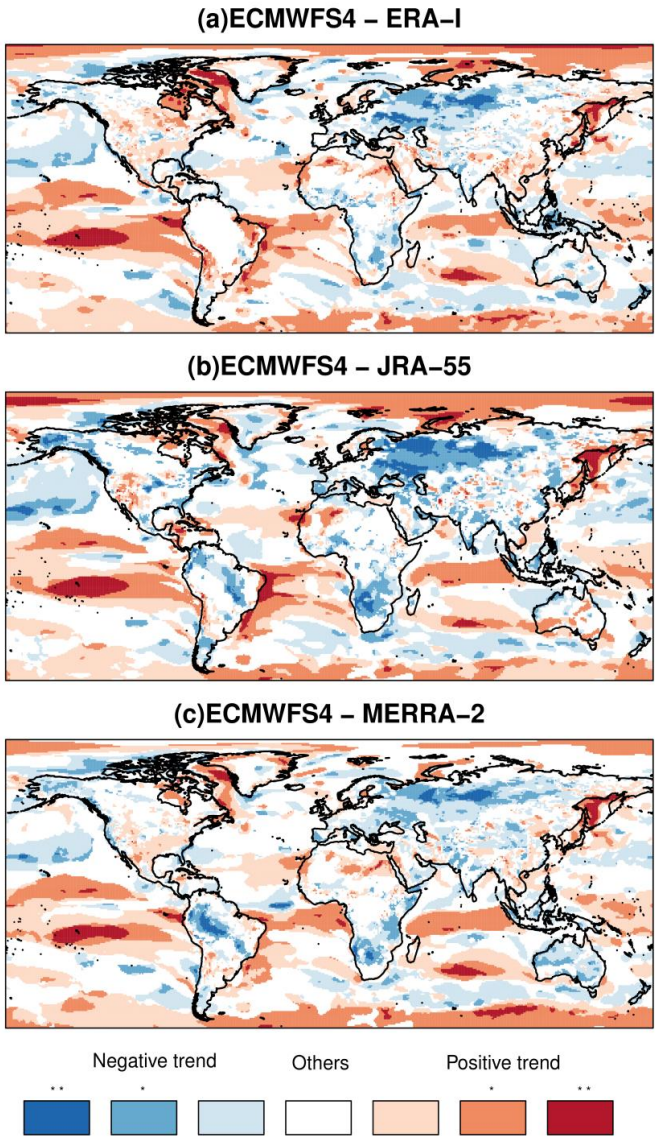
- Masking the results with a skill measure implies that we assume that the observations are perfect.
- Use of model output ignores the improvements that additional knowledge can provide to the user (e.g. using observed teleconnections).
- Significance tests used to decide where there is climate information do not tend to consider the power of the test (i.e. the probability of correctly rejecting the null hypothesis when it is false).
- No information of the physical processes is conveyed (e.g. forecast drift).
- Using a single forecast quality measure for the masking does not take into account its multifaceted complexity (skill, reliability).

Reference uncertainty: trends

(Bottom) Coherence of the 10-metre wind speed trends in three reanalyses (ERA-Interim, JRA-55 and MERRA) over 1981-2015 during boreal winter.

(Right) Coherence of the trends between ECMWF S4 and the three reanalyses.

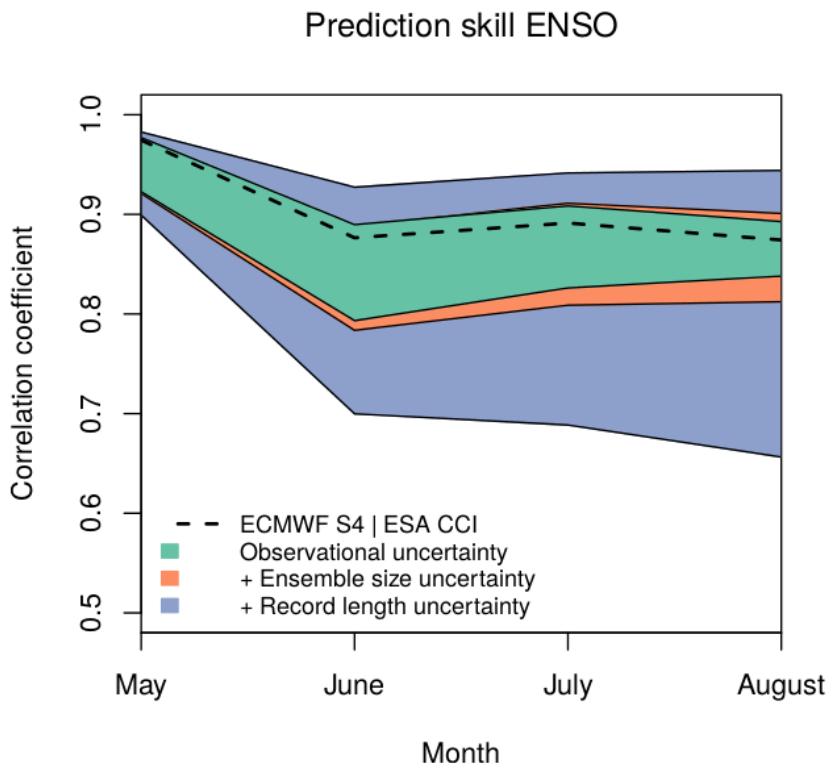


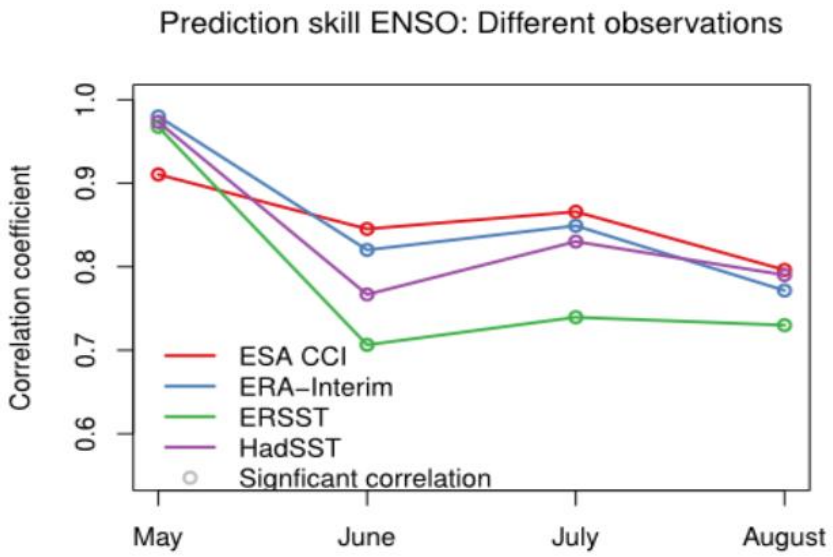


Torralba et al. (2017, ERL)

Sources of skill uncertainty

Niño 3.4 SST correlation of the ensemble mean for (left) EC-Earth3.1 (T511/ORCA025) predictions with ERAInt and GLORYS2v1 ics, and BSC sea-ice reconstruction and (right) ECMWF System 4, both started every May over 1993-2009.

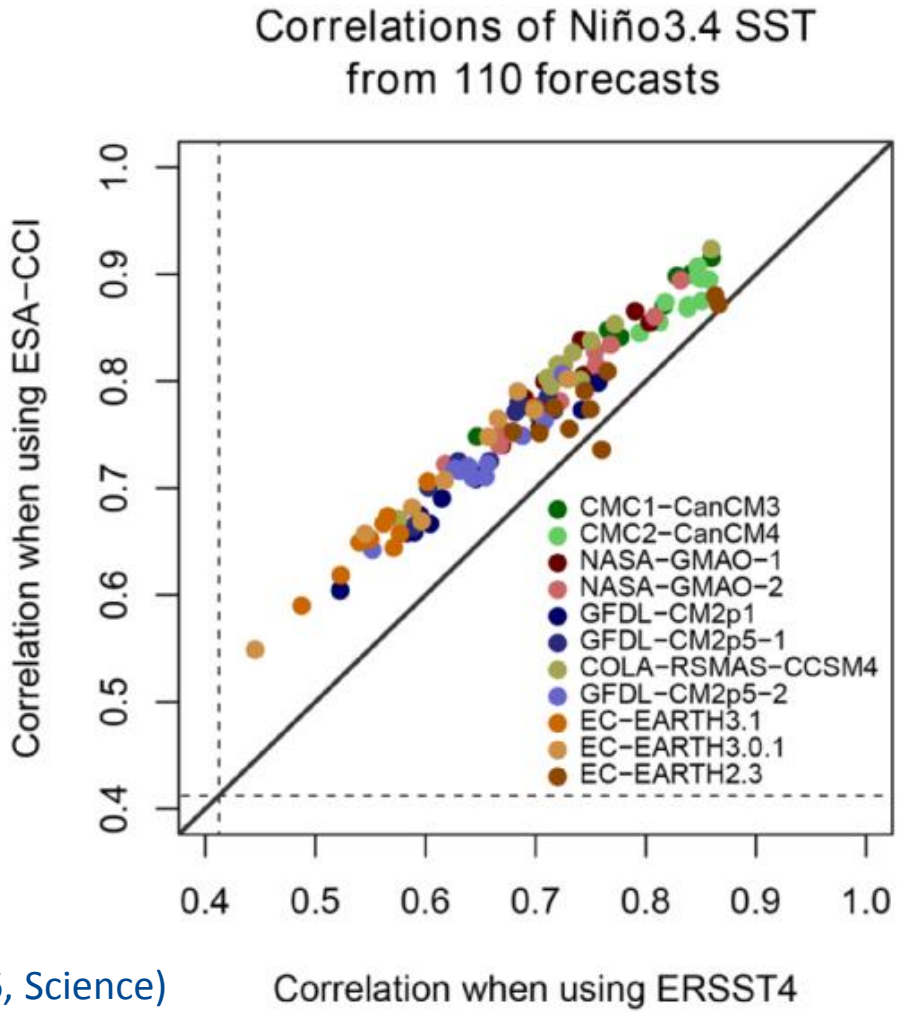




Bellprat et al. (2017, Rem. Sens. Env.)

Verifying with uncertain observations

Need to take into account the large observational uncertainty in the forecast quality estimates. Models can also be used to estimate the quality of observational estimates.



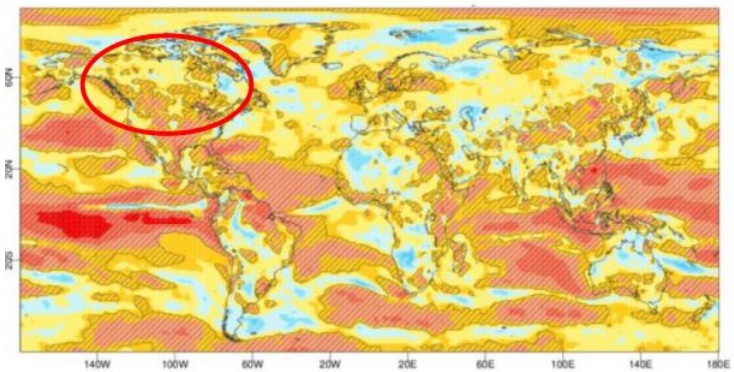
Massonnet et al. (2016, Science)

Using teleconnection information

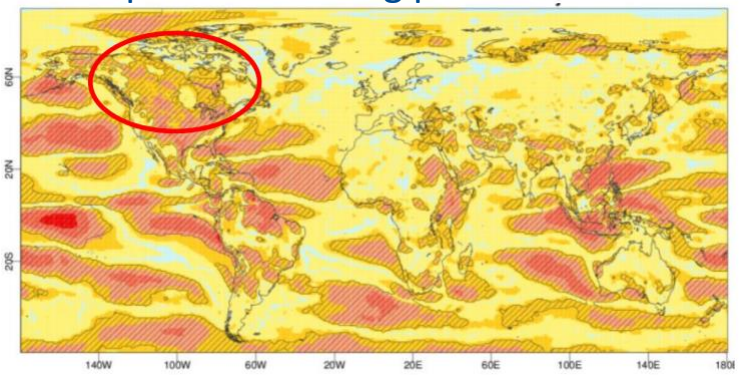
Barcelona Supercomputing Center
Centro Nacional de Supercomputación
EXCELENCIA SEVERO OCHOA

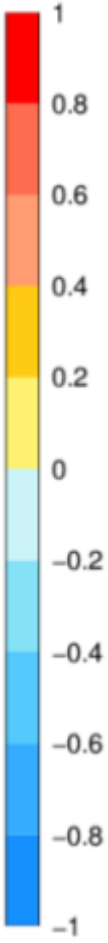
ECMWF S4 10-metre wind speed forecasts for DJF corrected with the predicted Niño3.4 index on a regression estimated using ERA-Interim.

Correlation of the ECMWF S4 ensemble-mean prediction (1981-2015)

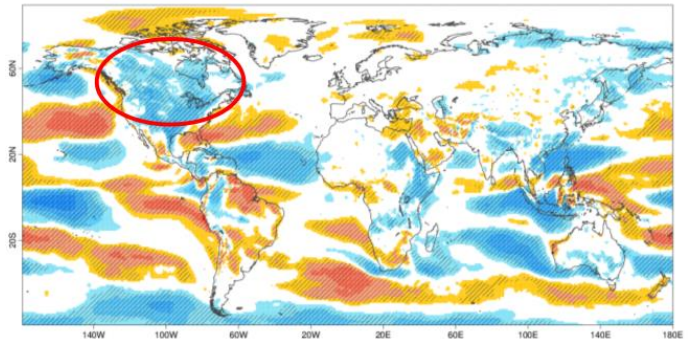


Correlation of the ECMWF S4 ensemble-mean prediction using predicted Niño3.4

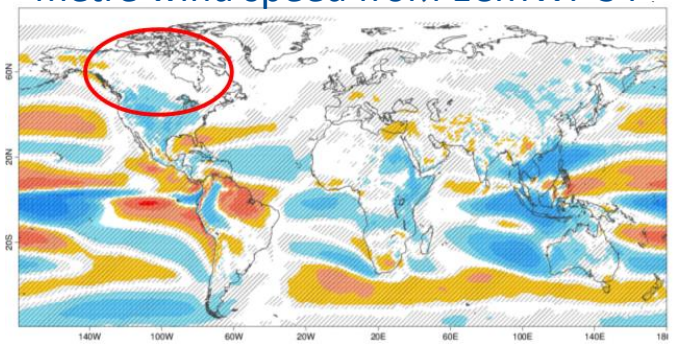


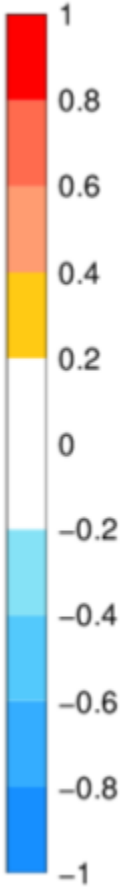


Point correlation of Niño3.4 and 10-metre wind speed from ERA Interim



Point correlation of Niño3.4 and 10-metre wind speed from ECMWF S4

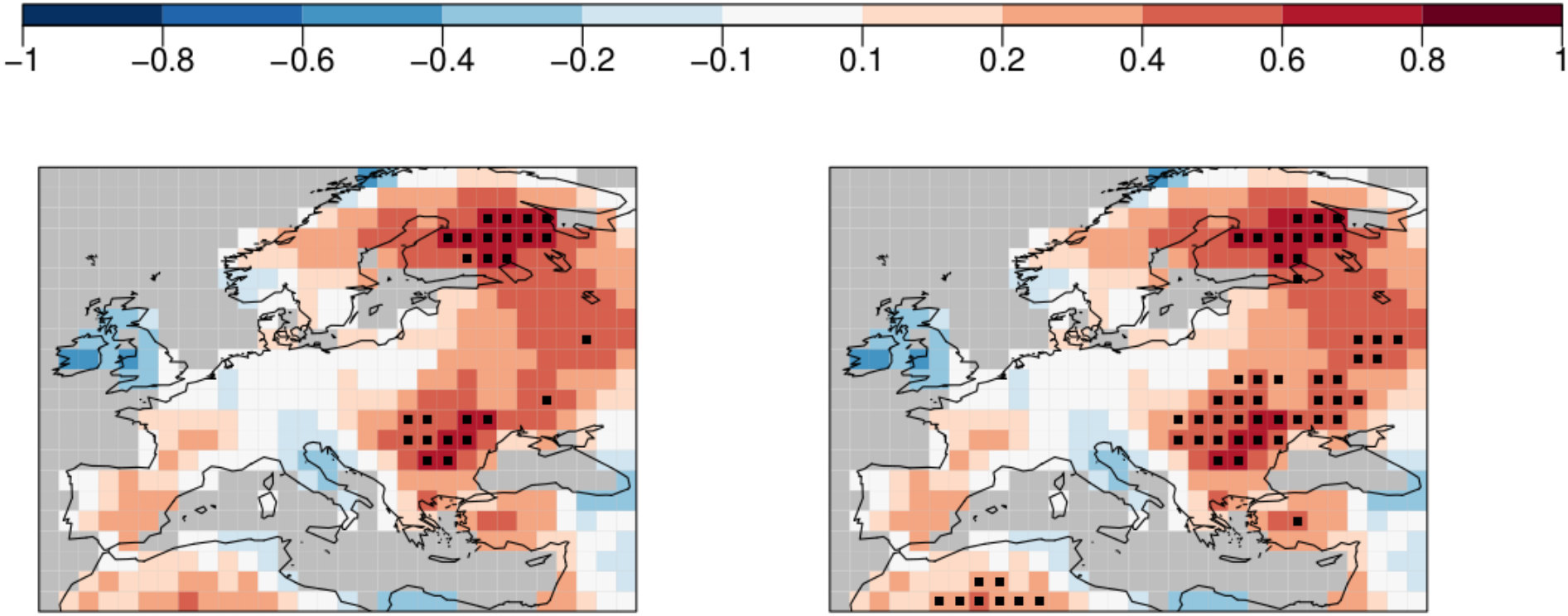




González et al. (in prep.)

Adequacy of the inference tests

Difference in correlation of EC-Earth3 (T511/ORCA025) seasonal one-month lead predictions started every May over 1993-2009 with ERAInt and GLORYS2v1 ics, and internal sea-ice reconstruction with realistic and climatological land-surface initial conditions: left for traditional Fisher test and right for Steiger test (with increased statistical power).

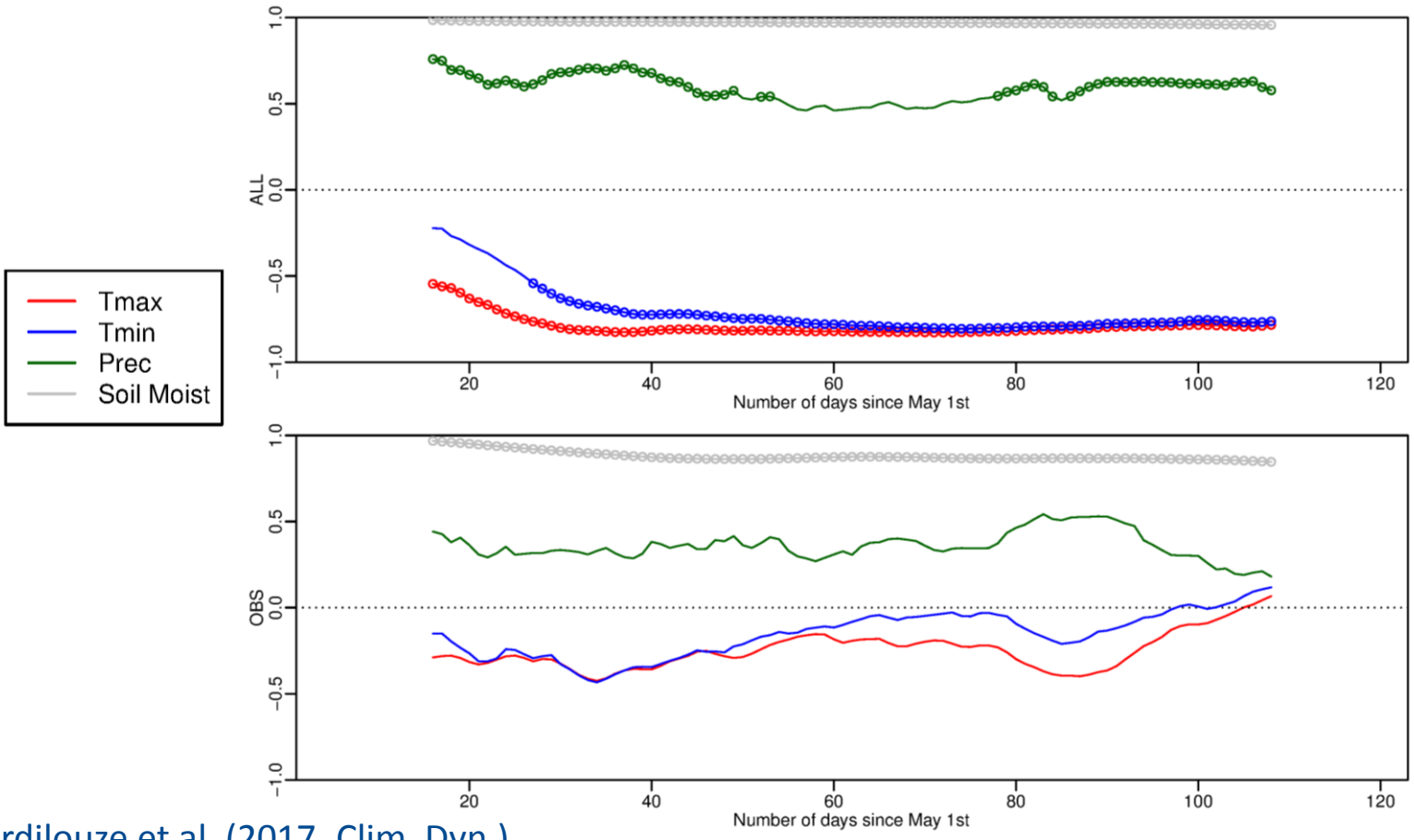


Siegert et al. (2016, MWR)

Drift helps uncovering model errors

Correlation between 1st of May total soil water content and 31-day running mean of variables from the SPECS multi-model seasonal forecast (top) and ERAInt (bottom) over North American Great Plains.

The model shifts quickly to excessive land-atmosphere coupling.



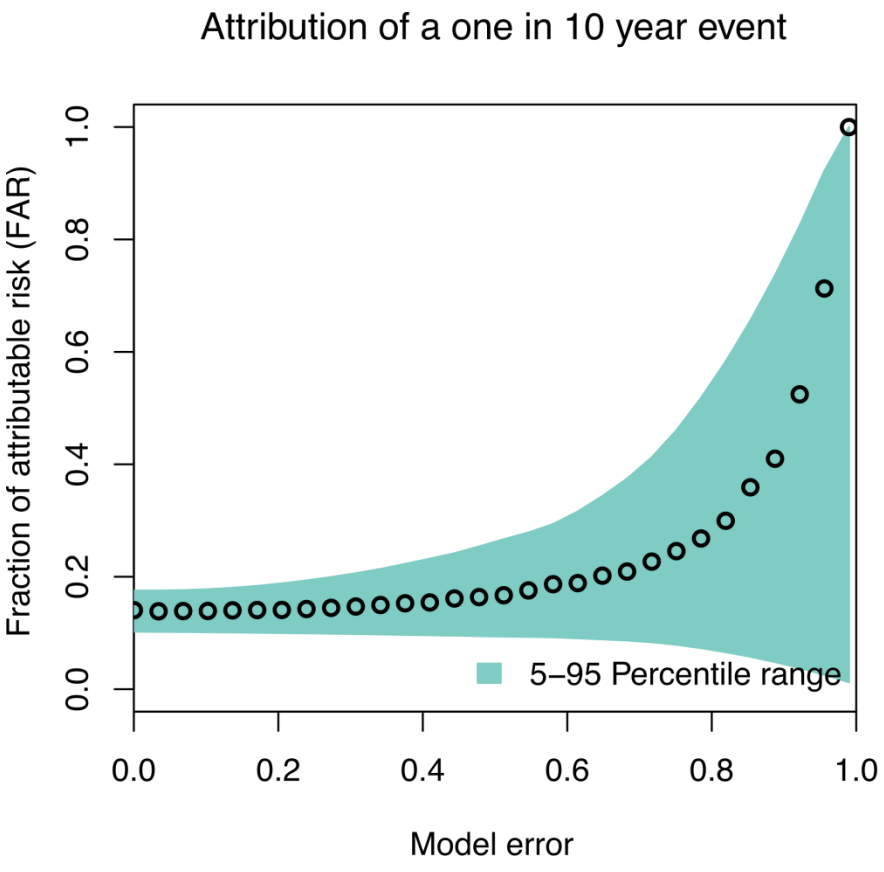
Ardilouze et al. (2017, Clim. Dyn.)

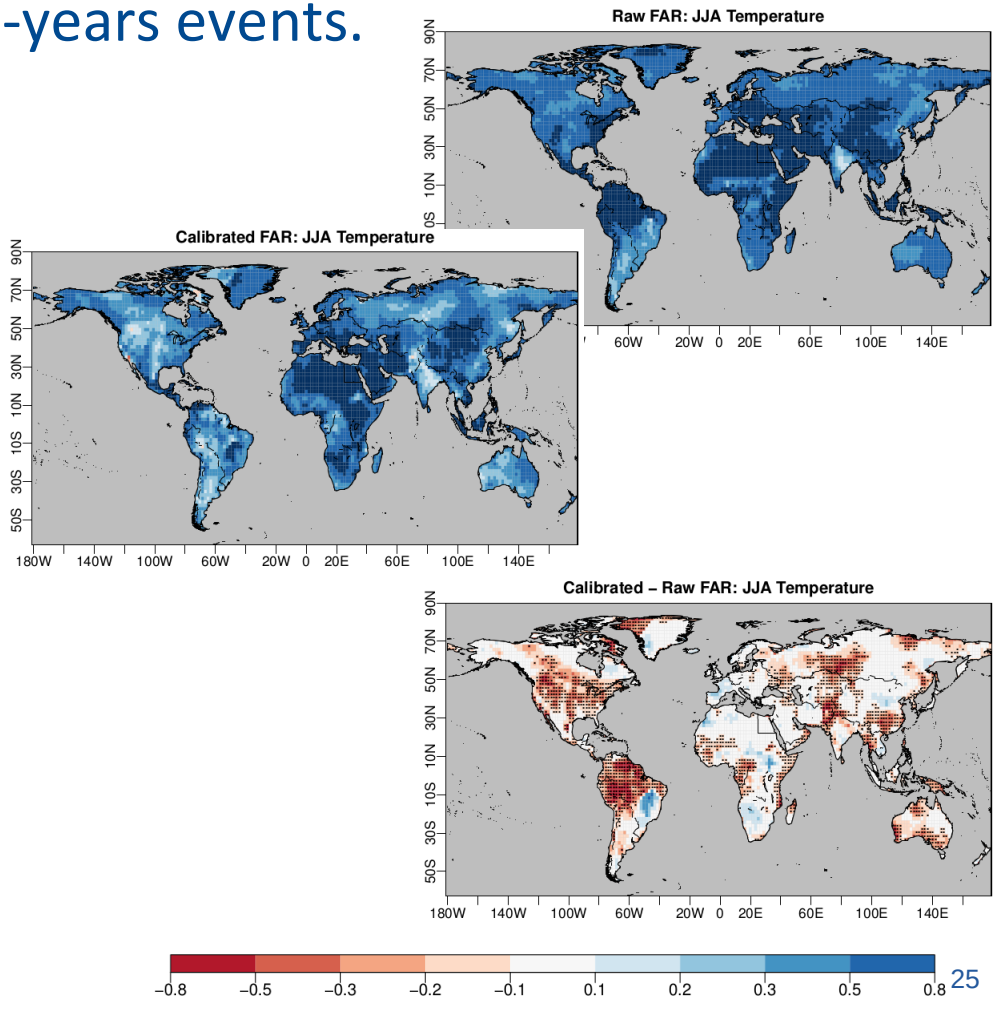
Verification: interpreting reliability

- Probability enables honest communication with the end user and making fair decisions.
- Reliability: if the probability of an event y is q the event should happen on average q 100% of the times P(y=1|p=q)=q.
- The reliability diagram is a plot of P(y=1|p) over p. Pointwise consistency bars by resampling and histogram of probabilities (sharpness diagram) should be added.
- The rank histogram verifies the raw ensemble, independent from the method used to obtain the probabilities, and measures whether the ensemble and the observation come from the same underlying distribution. It requires exchangeability and would benefit from testing for alternatives to flatness.

Attribution and reliability

Relationship between Fraction of Attributable Risk (FAR) and model error set in a toy model with reliability error. The FAR increases with the ensemble error. This is confirmed in a test with the HadGEM3A attribution system for one-in-ten-years events.





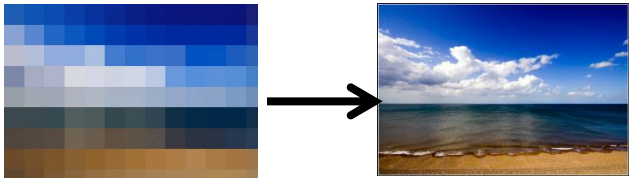
Bellprat and Doblas-Reyes (2016, GRL)

Providing better climate information

- Improve the physical processes in the models.
- Improve the initialisation and illustrate its impact.
- Better engage with the users.

Effect of increasing the resolution

Forecast quality from EC-Earth3.1 seasonal hindcasts (1993-2009, Glorys2v1, ERAInt and ERALand initial conditions). Solid for ESA-CCI and dashed for ERSST. Blue for high resolution ocean and atmosph[illegible]ed for high resolution ocean, black for standard resolution.

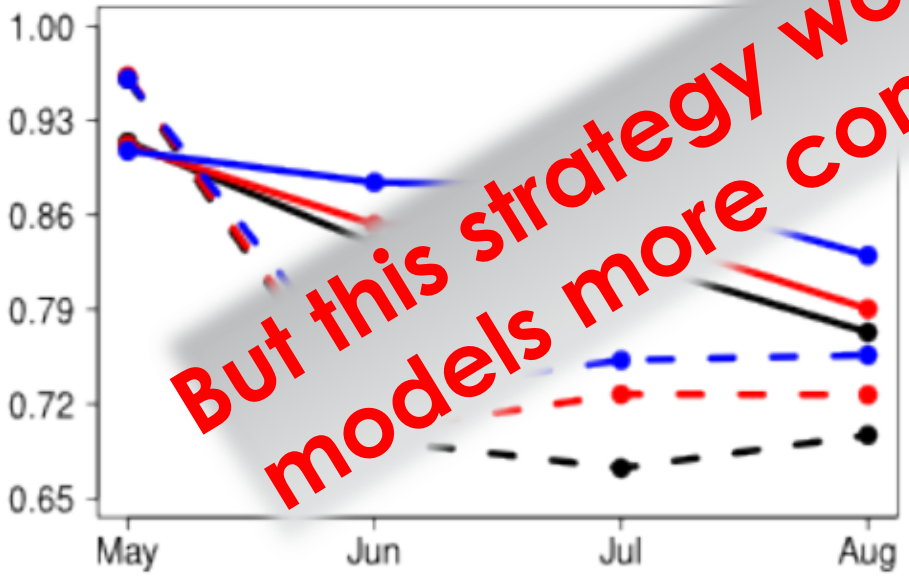


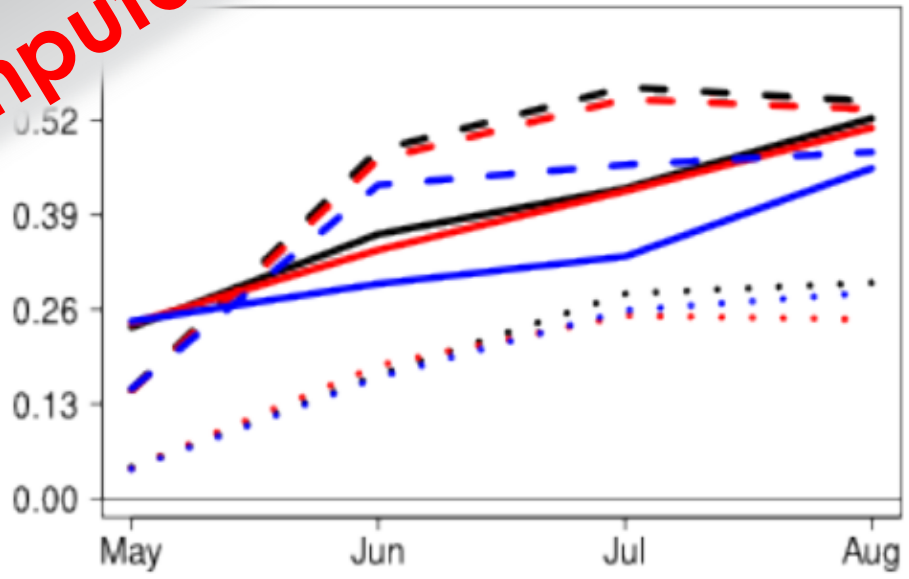
May start

SRes
IRes
HRes

a) Correlation

[illegible]ead and RMSE





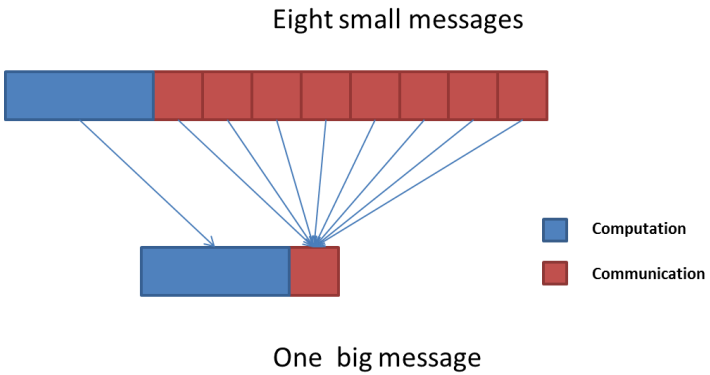
But this strategy works better making the models more computationally efficient

Prodhomme et al. (2016, J. Climate)

NEMO optimisation

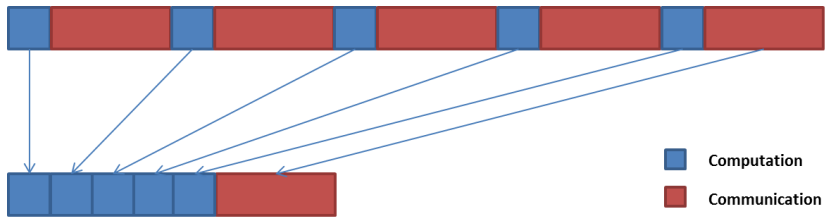
MPI message packing

Taking in account that NEMO is really sensitive to latency, messages aggregation is the best way to reduce the time invested in communications. Therefore, consecutive messages have been packed wherever the computational dependencies allow to do so.



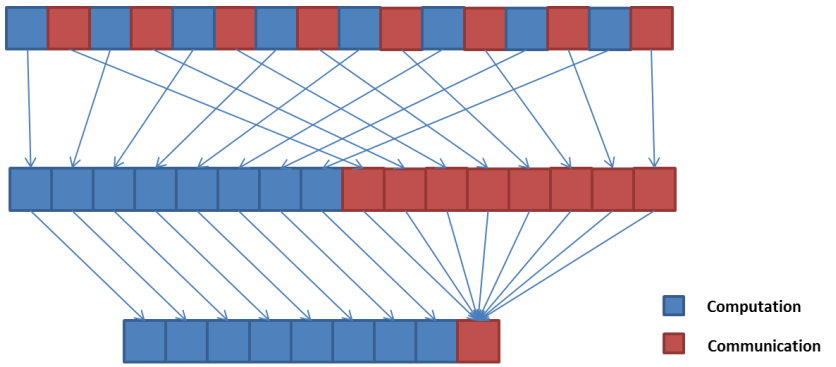
Convergence check reduction

Some routines use collective communications to perform a convergence check in iterative solvers. The cost of this verifications is really high, reaching a 66% of the time. Wherever the model allowed it, we reduced the frequency of this verifications in order to increase parallel efficiency.



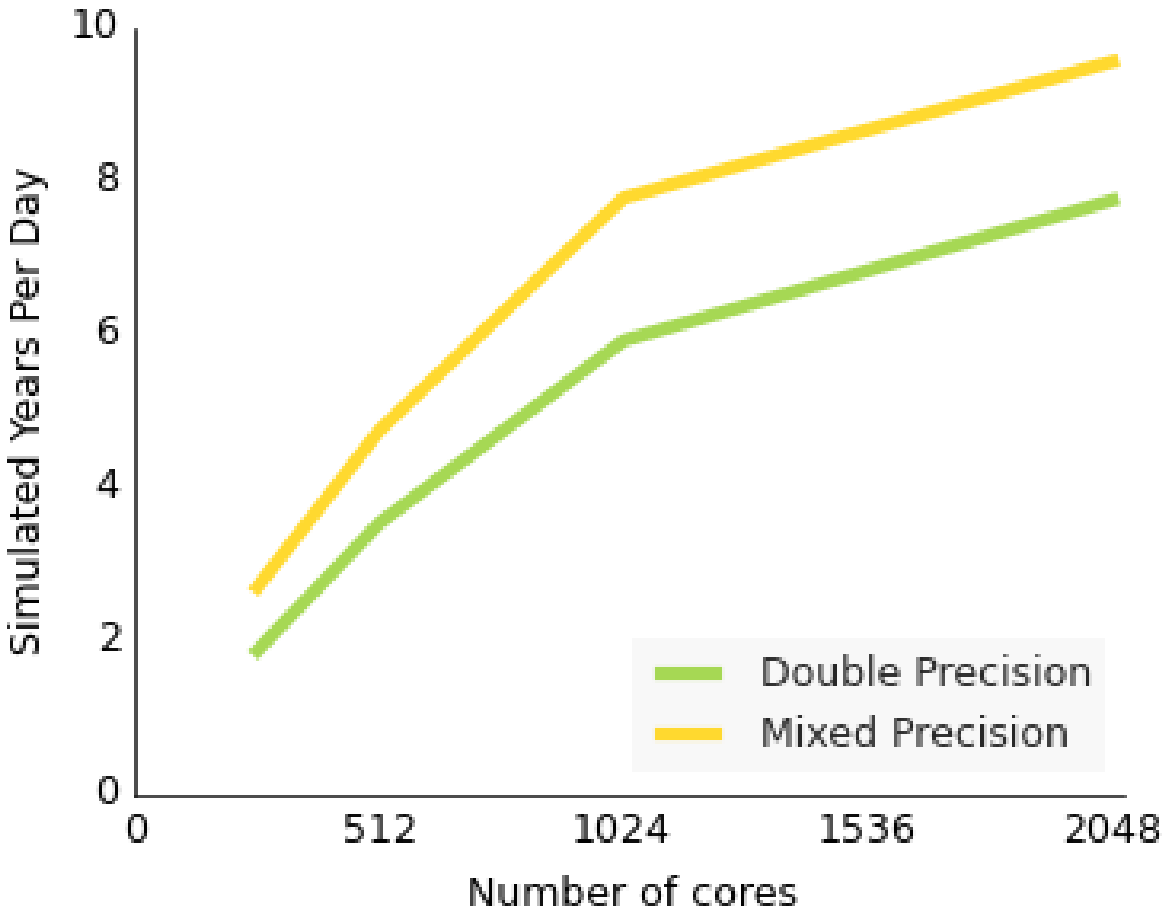
Reordering

In order to apply the message packing optimization to as many routines as possible, it was necessary to rearrange some computation and communication regions, taking into account the dependencies between them, to reduce the number of messages. This way it was possible to compute (and communicate) up to 41 variables at the same time, resulting in a dramatic reduction of the granularity.



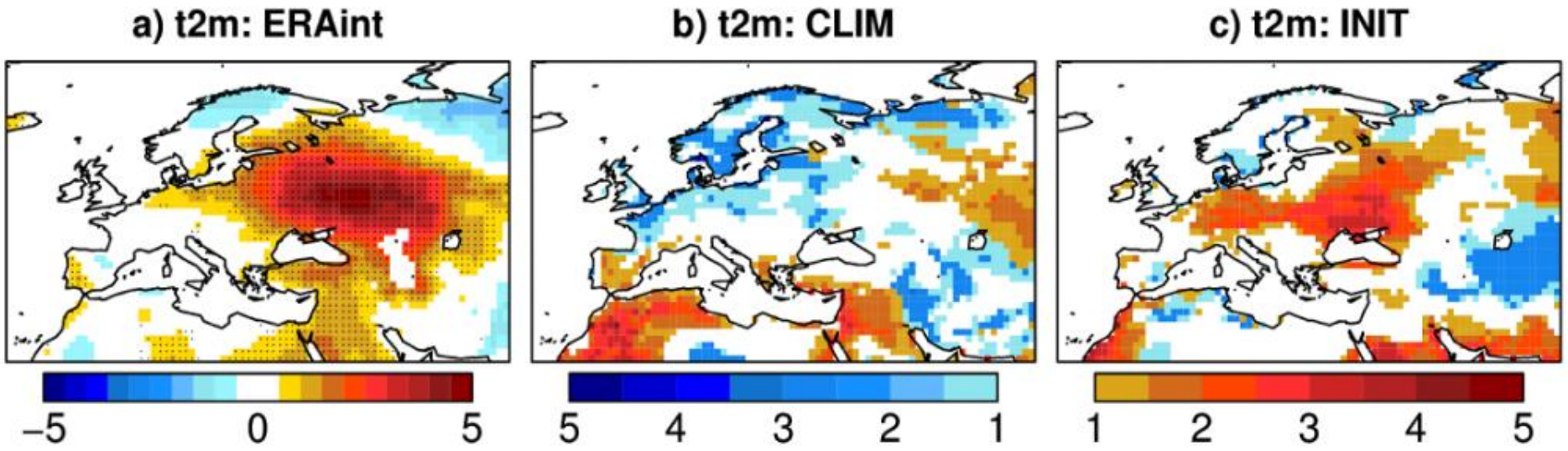
NEMO optimisation

Speed up of the NEMO3.6 (ORCA025L75) code when switching some parts of the code from double to single precision.



Improving initialisation

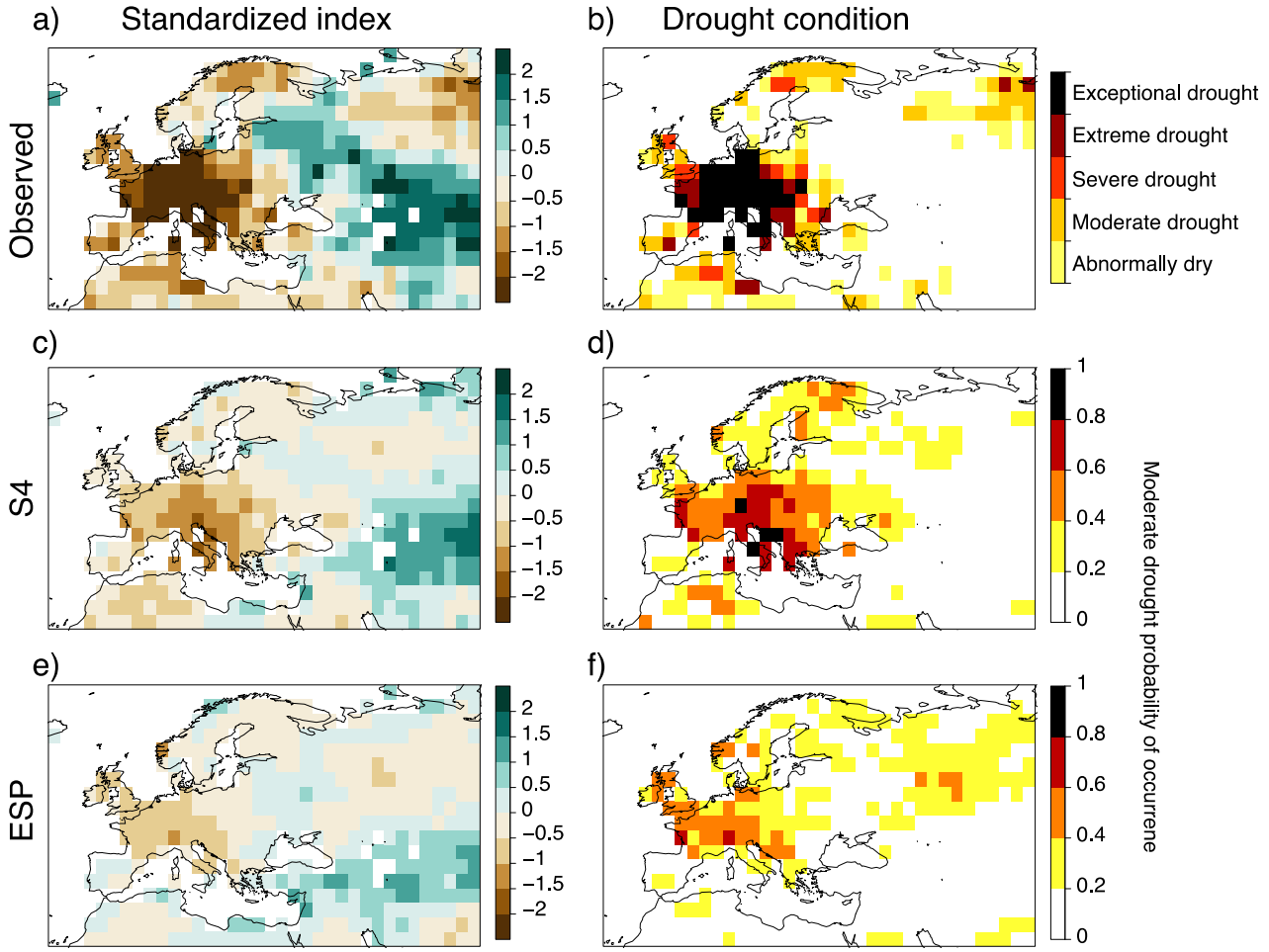
JJA near-surface temperature anomalies in 2010 from ERAInt (left) and odds ratio from experiments with a climatological (centre) and a realistic (right) land-surface initialisation. Results for EC-Earth2.3 started in May with initial conditions from ERAInt, ORAS4 and a sea-ice reconstruction over 1979-2010.



Prodhomme et al. (2015, Clim. Dyn.)

Check if the improvement matters

(Left) Six-month SPEI and (right) observed drought condition for August 2003 as observed, from ECMWF System 4 and from an ensemble streamflow prediction (ESP, from resampled historical data).



Turco et al. (submitted, ERL)

Predicting tropical storm frequency

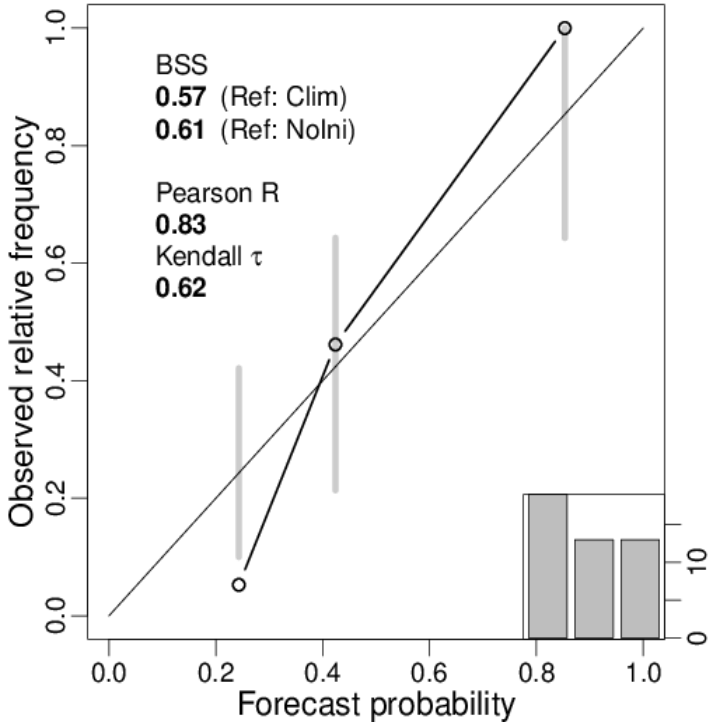
First comprehensive service of predictions of tropical cyclone seasonal frequency

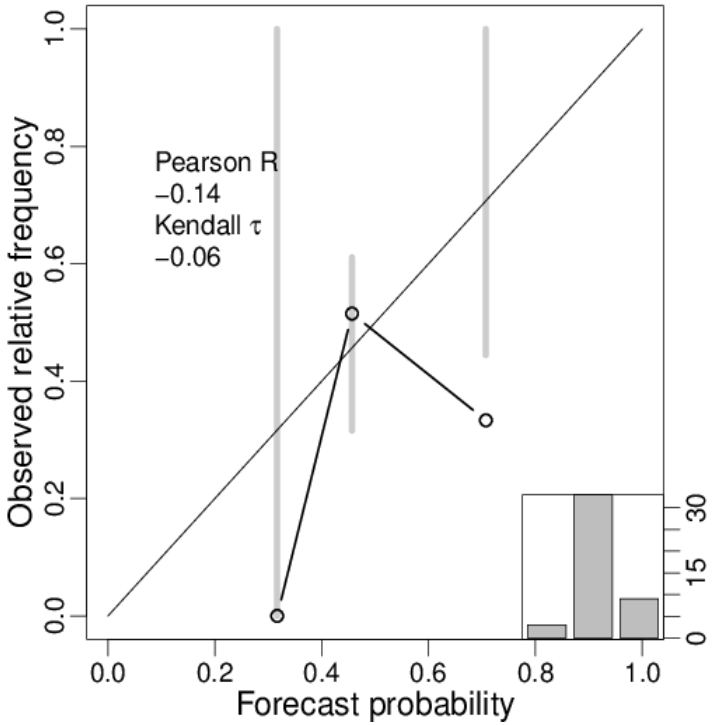
http://www.bsc.es/ESS/seasonalhurricanepredictions/

Predicting tropical storm frequency

Reliability diagrams of (left) initialised and (right) uninitialised MME simulations for basin-wide **accumulated cyclone energy** (ACE). Results are for 2-9 year means above the climatological median from decadal predictions over 1961-2009. Statistically significant values in bold.

Some of the added value of the predictions is their better management of uncertainty, which leads to increased **credibility**.

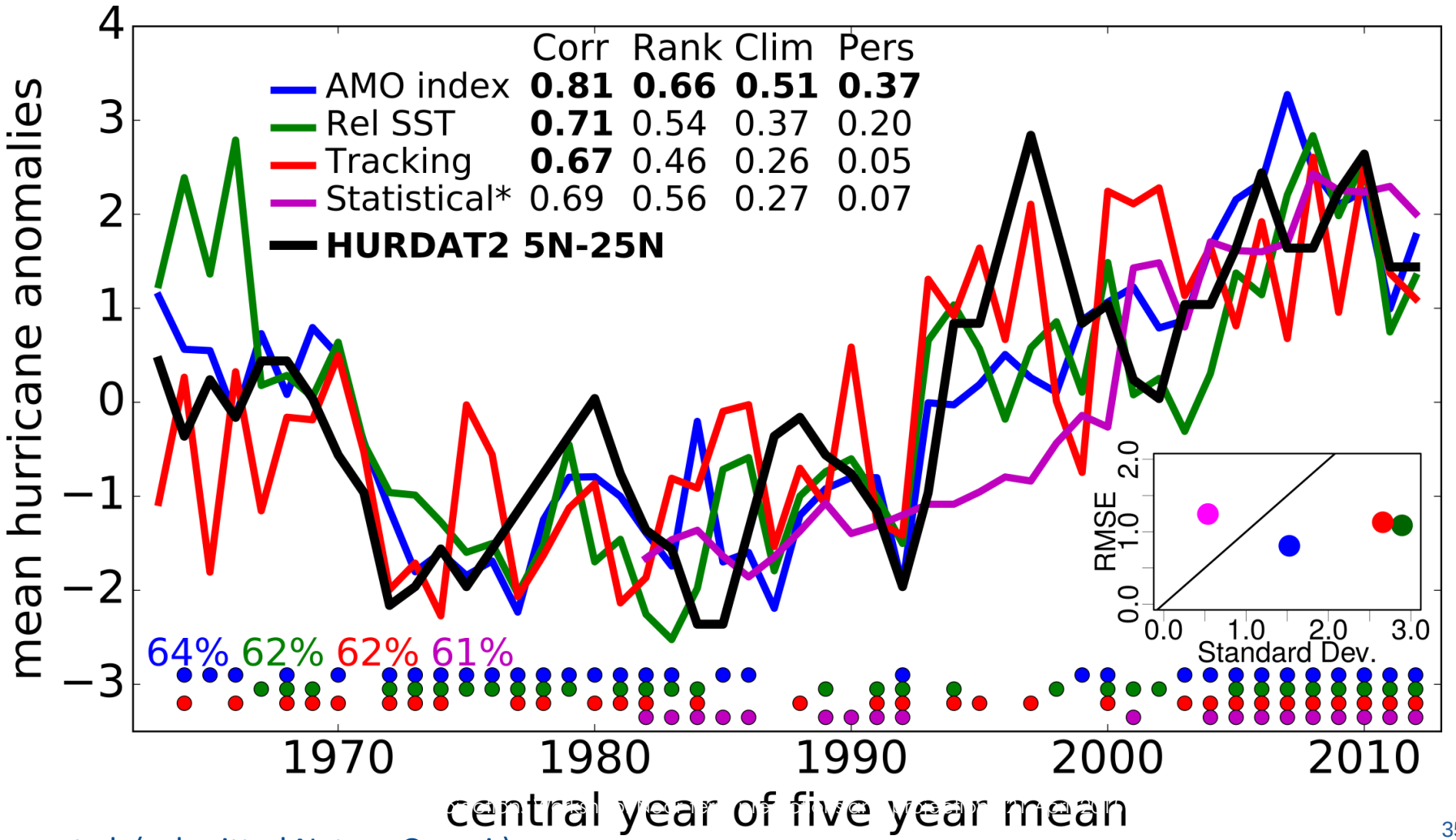




Caron et al. (2015, GRL)

Service-driven predictions

Multi-model decadal predictions of 1-5 year mean hurricane number per year using different methods.



Caron et al. (submitted Nature Geosci.)

Some ideas: research and services

A new paradigm has come to stay: user-driven research

- **Progress**: opportunities appear in a context where research and services are closer together.
- **Education**: in the era of open data, take advantage of the open education opportunities.
- **Heterogeneity**: link to and merge our data with communities with larger impact (urban, arts, social).
- **Standards**: in a collaborative environment standards are a must and everyone's responsibility.
- **Technology**: make the most of a rapidly evolving technology (heterogeneous nodes, software, mobile data capture, visualisation, storage/compression, computing and storage outsourcing).
- **Industry engagement:** how can we involve the private sector more efficiently?